Figure 1

M-L-S-N-L-R-I-L-L-N-K-A-A-L-R-K-A-H-T-S-M-V-R-N-F-R-Y-G-K-P-V-Q-S-Q-L-K-P-R-D-L-C

amino-terminal end ⟷ carboxy-terminal end

22

CCCCGGGTACCTTGCGAGCCCTGGGCTCGCAAGGTACCC

5´-end ⟷ 3´-end

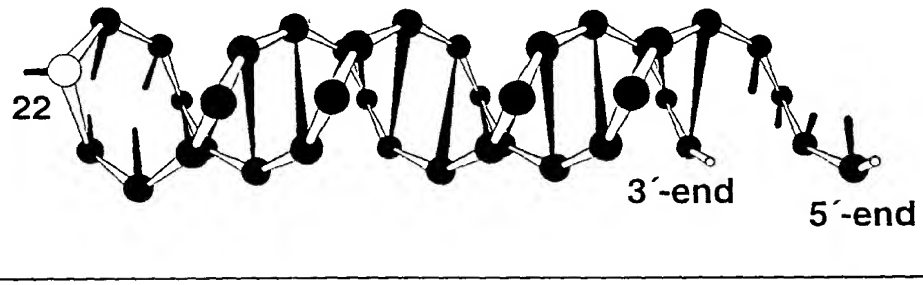


Figure 2

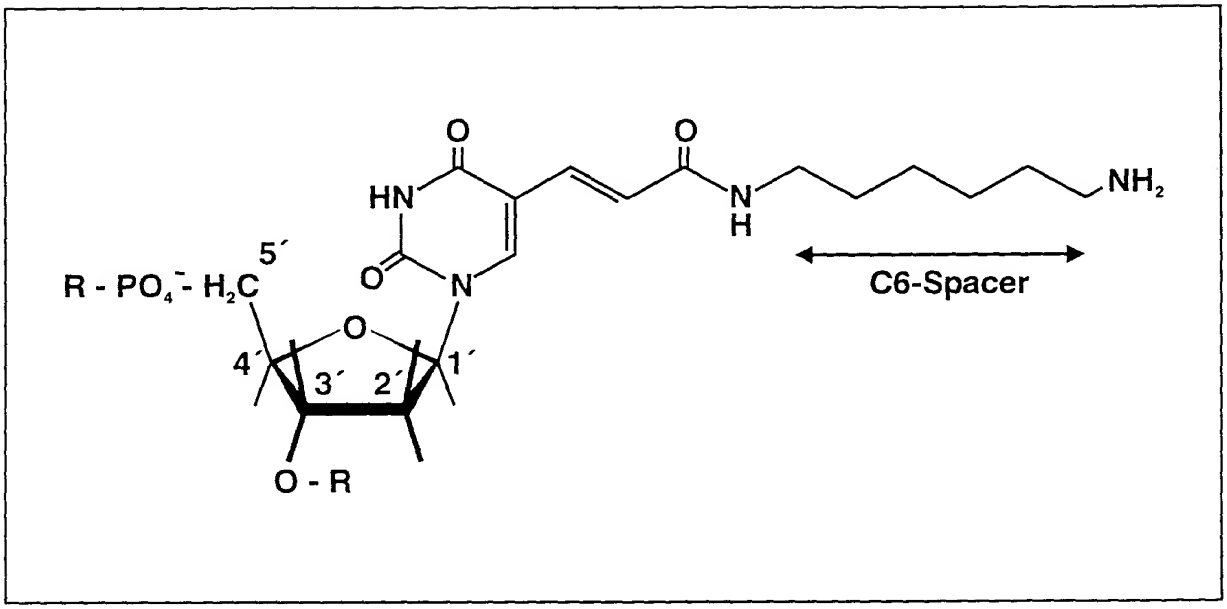


Figure 3

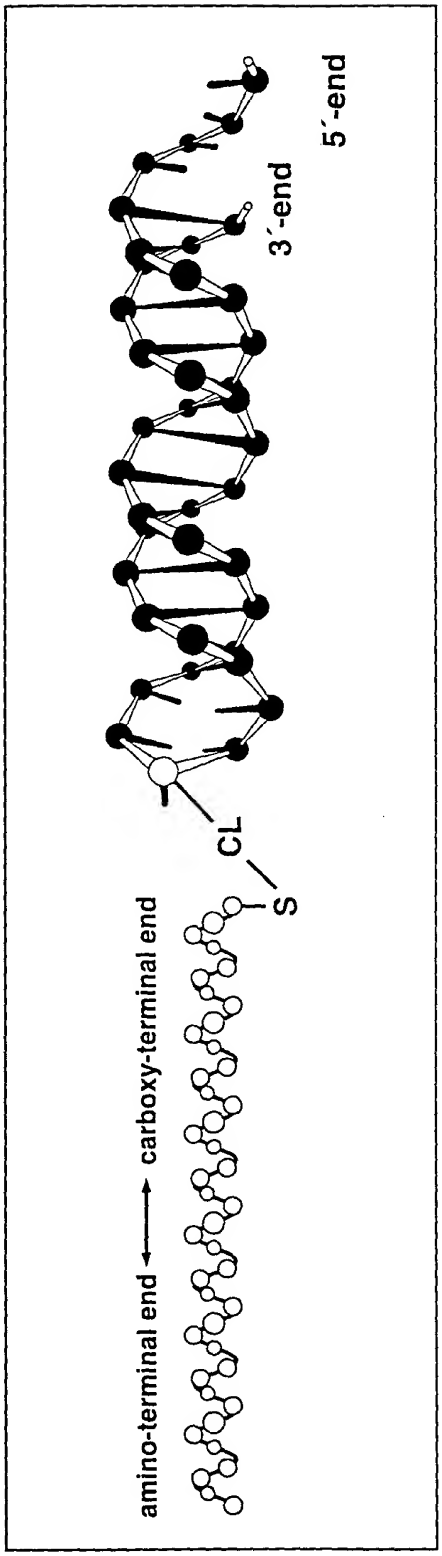


Figure 4

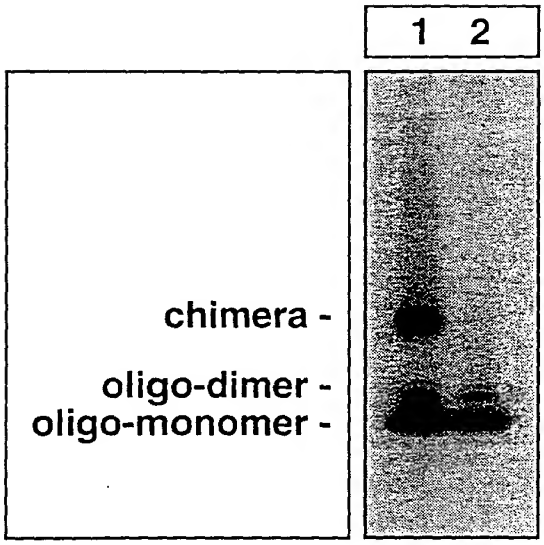


Figure 5a

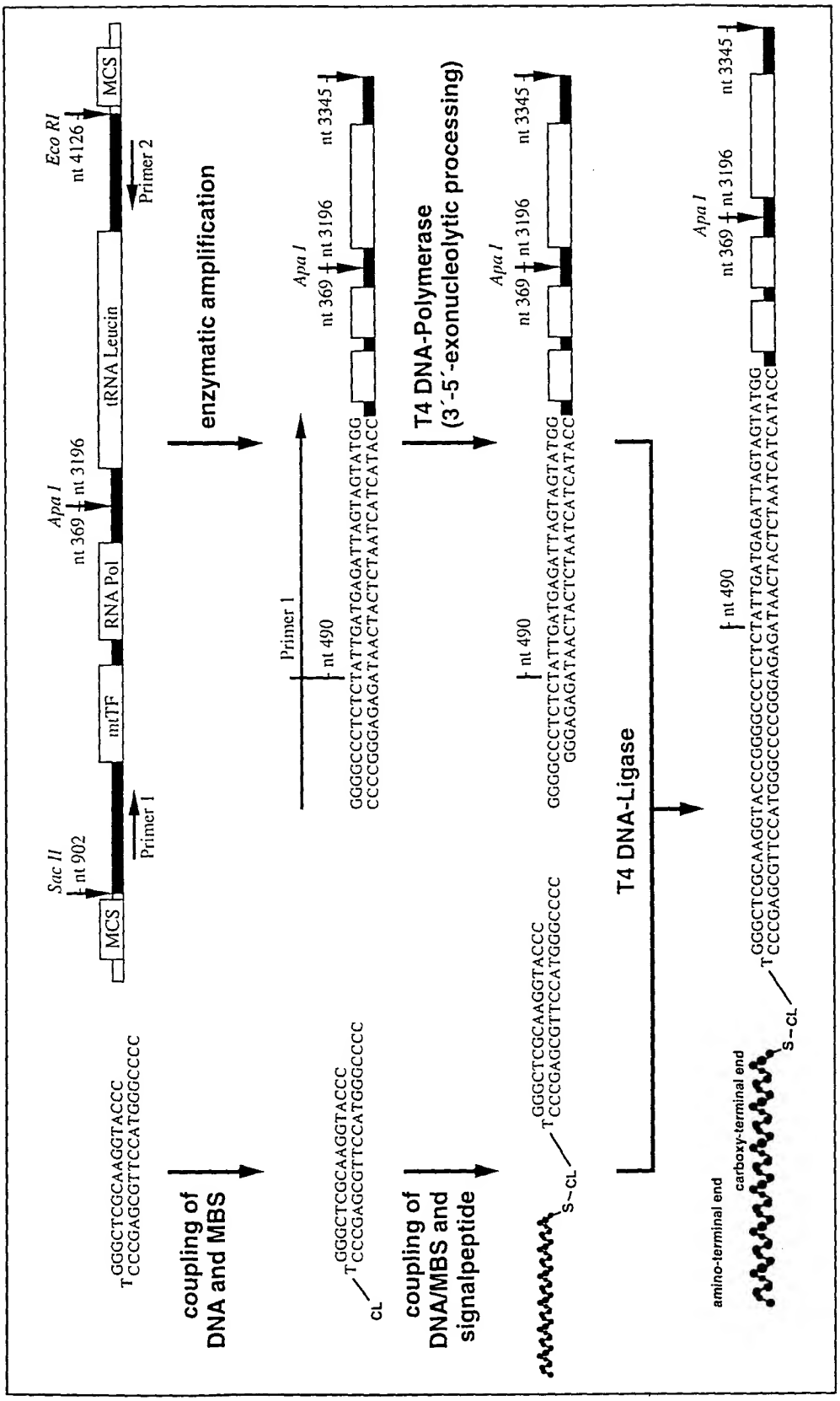

Sac II
nt 902
MCS
Primer 1
Apa I
nt 369
nt 3196
mTF
RNA Pol
tRNA Leucin
Eco RI
nt 4126
MCS
Primer 2
enzymatic amplification
Primer 1
nt 490
Apa I
nt 369
nt 3196
nt 3345
T4 DNA-Polymerase
(3´-5´-exonucleolytic processing)
T4 DNA-Ligase
coupling of DNA and MBS
coupling of DNA/MBS and signalpeptide
S-CL
amino-terminal end
carboxy-terminal end

Figure 5b

```
        10          20          30          40          50          60
CCGCGGTGGC TGGCACGAAA TTGACCAACC CTGGGGTTAG TATAGCTTAG TAAACTTTC
GGCGCCACCG ACCGTGCTTT AACTGGTTGG GACCCCAATC ATATCGAATC ATTTGAAAG

        70          80          90          100         110         120
GTTTATTGCT AAAGGTTAAT CACTGCTGTT TCCCGTGGGG GTGTGGCTAG CTAAGCGTT
CAAATAACGA TTTCCAATTA GTGACGACAA AGGGCACCCC CACACCGATC GATTCGCAA

        130         140         150         160         170         180
TTGAGCTGCA TTGCTGCGTG CTTGATGCTT GTTCCTTTTG ATCGTGGTGA TTAGAGGGT
AACTCGACGT AACGACGCAC GAACTACGAA CAAGGAAAAC TAGCACCACT AATCTCCCA

        190         200         210         220         230         240
GAACTCACTG GAACGGGGAT GCTTGCATGT GTAATCTTAC TAAGAGCTAA AGAAAGGCT
CTTGAGTGAC CTTGCCCCTA CGAACGTACA CATTAGAATG ATTCTCGATT TCTTTCCGA

        250         260         270         280         290         300
AGGACCAAAC CTATTTGTTT ATGGGGTGAT GTGAGCCCGT CTAAACATTT CAGTGTATT
TCCTGGTTTG GATAAACAAA TACCCCACTA CACTCGGGCA GATTTGTAAA GTCACATAA

        310         320         330         340         350         360
GCTTTGAGGA GGTAAGCTAC ATAAACTGTG GGGGGTGTCT TTGGGGTTTG TTGGTTCGG
CGAAACTCCT CCATTCGATG TATTTGACAC CCCCCACAGA AACCCCAAAC AACCAAGCC

        370         380         390         400         410         420
GGTATGGGGT TAGCAGCGGT GTGTGTGTGC TGGGTAGGAT GGGCGGGGGT GTATTGATG
CCATACCCCA ATCGTCGCCA CACACACACG ACCCATCCTA CCCGCCCCCA CATAACTAC

        430         440         450         460         470         480
AGATTAGTAG TATGGGAGTG GGAGGGGAAA ATAATGTGTT AGTTGGGGGG GACTGTTAA
TCTAATCATC ATACCCTCAC CCTCCCCTTT TATTACACAA TCAACCCCCC CTGACAATT

        490         500         510         520         530         540
AAGTGCATAC CGCCAAAAGA TAAAATTTGA AATCTGGTTA GGCTGGTGTT GGGCCCTTT
TTCACGTATG GCGGTTTTCT ATTTTAAACT TTAGACCAAT CCGACCACAA CCCGGGAAA

        550         560         570         580         590         600
GTCCCACACC CACCCAAGAA CAGGGTTTGT TAAGATGGCA GAGCCCGGTA TCGCATAAA
CAGGGTGTGG GTGGGTTCTT GTCCCAAACA ATTCTACCGT CTCGGGCCAT AGCGTATTT

        610         620         630         640         650         660
ACTTAAAACT TTACAGTCAG AGGTTCAATT CCTCTTCTTA ACAACATACC ATGGCCAAC
TGAATTTTGA AATGTCAGTC TCCAAGTTAA GGAGAAGAAT TGTTGTATGG TACCGGTTG

        670         680         690         700         710         720
CTCCTACTCC TCATTGTACC CATTCTAATC GCAATGGCAT TCCTAATGCT ACCGAACGA
GAGGATGAGG AGTAACATGG GTAAGATTAG CGTTACCGTA AGGATTACGA TGGCTTGCT

        730         740         750         760         770         780
AAAATTCTAG GCTATATACA ACTACGCAAA GGCCCCAACG TGGTAGGCCC TACGGGCTA
TTTTAAGATC CGATATATGT TGATGCGTTT CCGGGGTTGC ACCATCCGGG ATGCCCGAT

        790         800         810         820         830         840
CTACAACCCT TCGCTGACGC CATAAAACTC TTCACCAAAG AGCCCCTAAA CCCGCCACA
GATGTTGGGA AGCGACTGCG GTATTTTGAG AAGTGGTTTC TCGGGGATTT GGGCGGTGT
```

```
       850         860         870         880         890         900
TCTACCATCA CCCTCTACAT CACCGCCCCG ACCTTAGCTC TCACCATCGC CTTCTACTA
AGATGGTAGT GGGAGATGTA GTGGCGGGGC TGGAATCGAG AGTGGTAGCG GAAGATGAT

       910         920         930         940         950         960
TGAACCCCCC TCCCCATACC CAACCCCCTG GTCAACCTCA ACCTAGGCCT CTATTTATT
ACTTGGGGGG AGGGGTATGG GTTGGGGGAC CAGTTGGAGT TGGATCCGGA GATAAATAA

       970         980         990        1000        1010        1020
CTAGCCACCT CTAGCCTAGC CGTTTACTCA ATCCTCTGAT CAGGGTGAGC TCAAACTCA
GATCGGTGGA GATCGGATCG GCAAATGAGT TAGGAGACTA GTCCCACTCG AGTTTGAGT

      1030        1040        1050        1060        1070        1080
AACTACGCCC TGATCGGCGC ACTGCGAGCA GTAGCCCAAA CAATCTCATA GAAGTCACC
TTGATGCGGG ACTAGCCGCG TGACGCTCGT CATCGGGTTT GTTAGAGTAT CTTCAGTGG

      1090        1100        1110        1120        1130        1140
CTAGCCATCA TTCTACTATC AACATTACTA ATAAGTGGCT CCTTTAACCT TCCACCCTT
GATCGGTAGT AAGATGATAG TTGTAATGAT TATTCACCGA GGAAATTGGA AGGTGGGAA

      1150        1160        1170        1180        1190        1200
ATCACAACAC AAGAACACCT CTGATTACTC CTGCCATCAT GACCCTTGGC ATAATATGA
TAGTGTTGTG TTCTTGTGGA GACTAATGAG GACGGTAGTA CTGGGAACCG TATTATACT

      1210        1220        1230        1240        1250        1260
TTTATCTCCA CACTAGCAGA GACCAACCGA ACCCCCTTCG ACCTTGCCGA GGGGAGTCC
AAATAGAGGT GTGATCGTCT CTGGTTGGCT TGGGGGAAGC TGGAACGGCT CCCCTCAGG

      1270        1280        1290        1300        1310        1320
GAACTAGTCT CAGGCTTCAA CATCGAATAC GCCGCAGGCC CCTTCGCCCT TTCTTCATA
CTTGATCAGA GTCCGAAGTT GTAGCTTATG CGGCGTCCGG GGAAGCGGGA AAGAAGTAT

      1330        1340        1350        1360        1370        1380
GCCGAATACA CAAACATTAT TATAATAAAC ACCCTCACCA CTACAATCTT CTAGGAACA
CGGCTTATGT GTTTGTAATA ATATTATTTG TGGGAGTGGT GATGTTAGAA GATCCTTGT

      1390        1400        1410        1420        1430        1440
ACATATGACG CACTCTCCCC TGAACTCTAC ACAACATATT TTGTCACCAA ACCCTACTT
TGTATACTGC GTGAGAGGGG ACTTGAGATG TGTTGTATAA AACAGTGGTT TGGGATGAA

      1450        1460
CTAACCTCCC TGTTCTTATG AATTC
GATTGGAGGG ACAAGAATAC TTAAG
```

Figure 6a

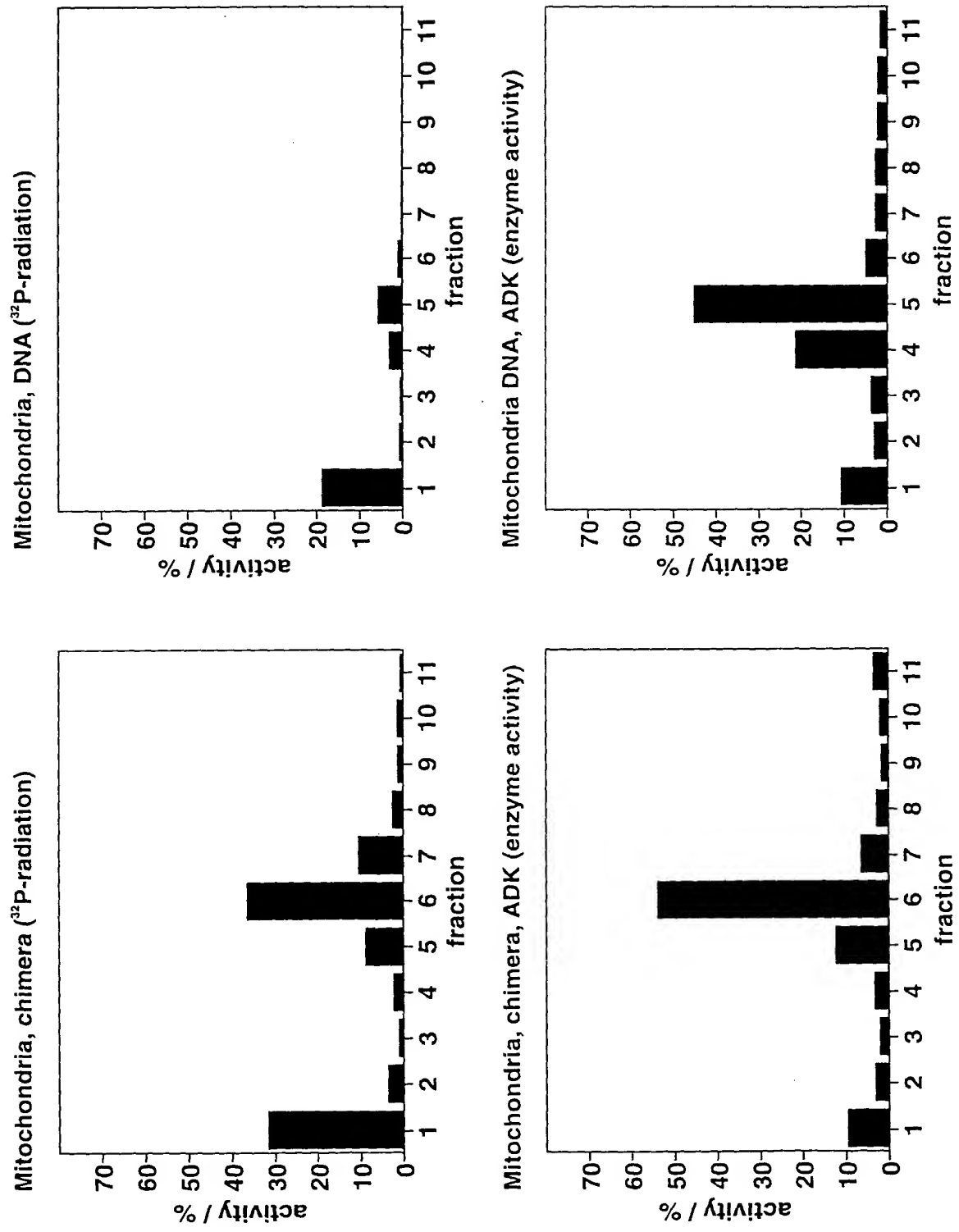

Mitochondria, DNA (^{32}P-radiation)
Mitochondria DNA, ADK (enzyme activity)
Mitochondria, chimera (^{32}P-radiation)
Mitochondria, chimera, ADK (enzyme activity)
activity / %
fraction
0
10
20
30
40
50
60
70
1 2 3 4 5 6 7 8 9 10 11

Figure 6b

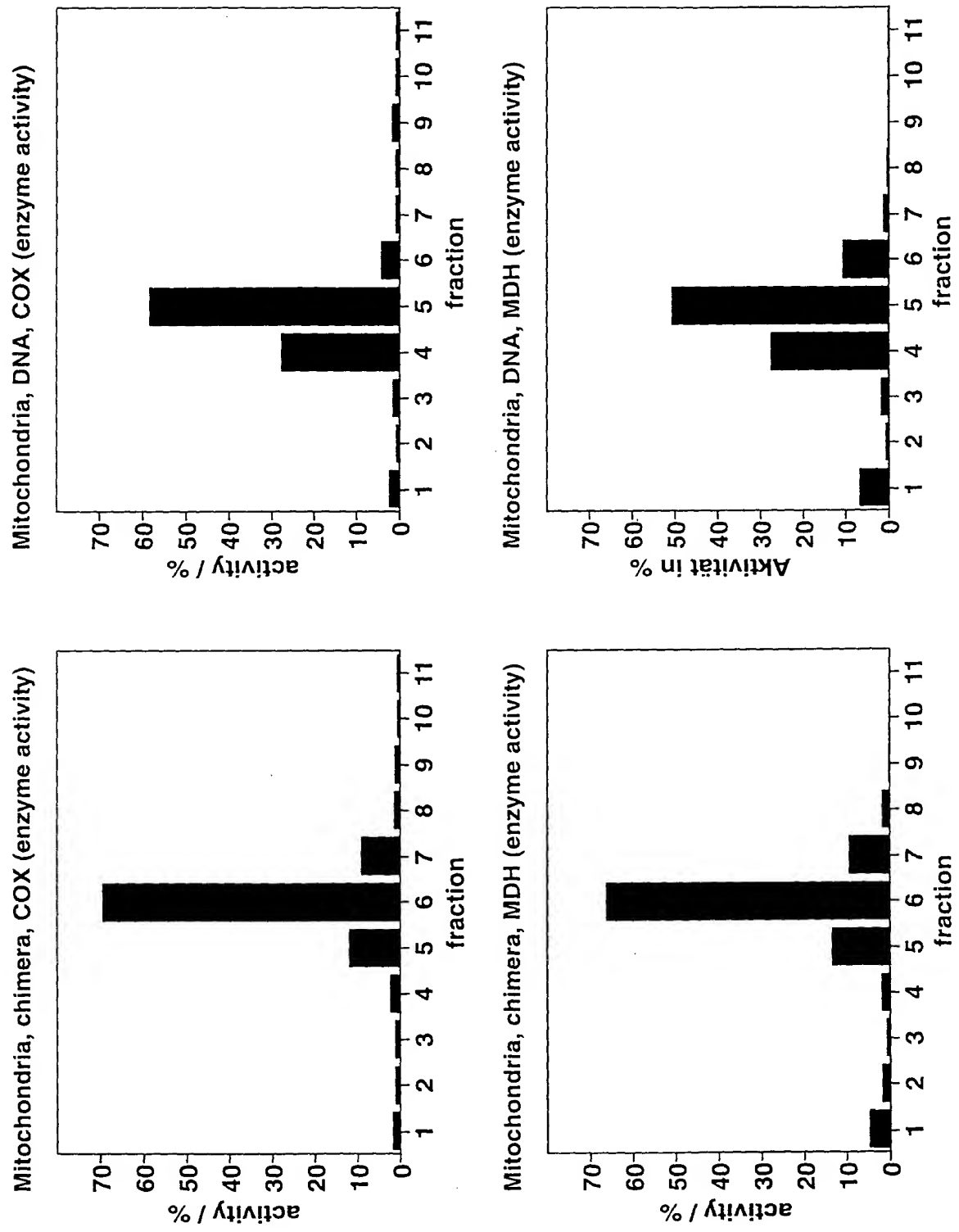

Mitochondria, DNA, COX (enzyme activity)
activity / %
fraction
Mitochondria, DNA, MDH (enzyme activity)
Aktivität in %
fraction
Mitochondria, chimera, COX (enzyme activity)
activity / %
fraction
Mitochondria, chimera, MDH (enzyme activity)
activity / %
fraction

Figure 7a

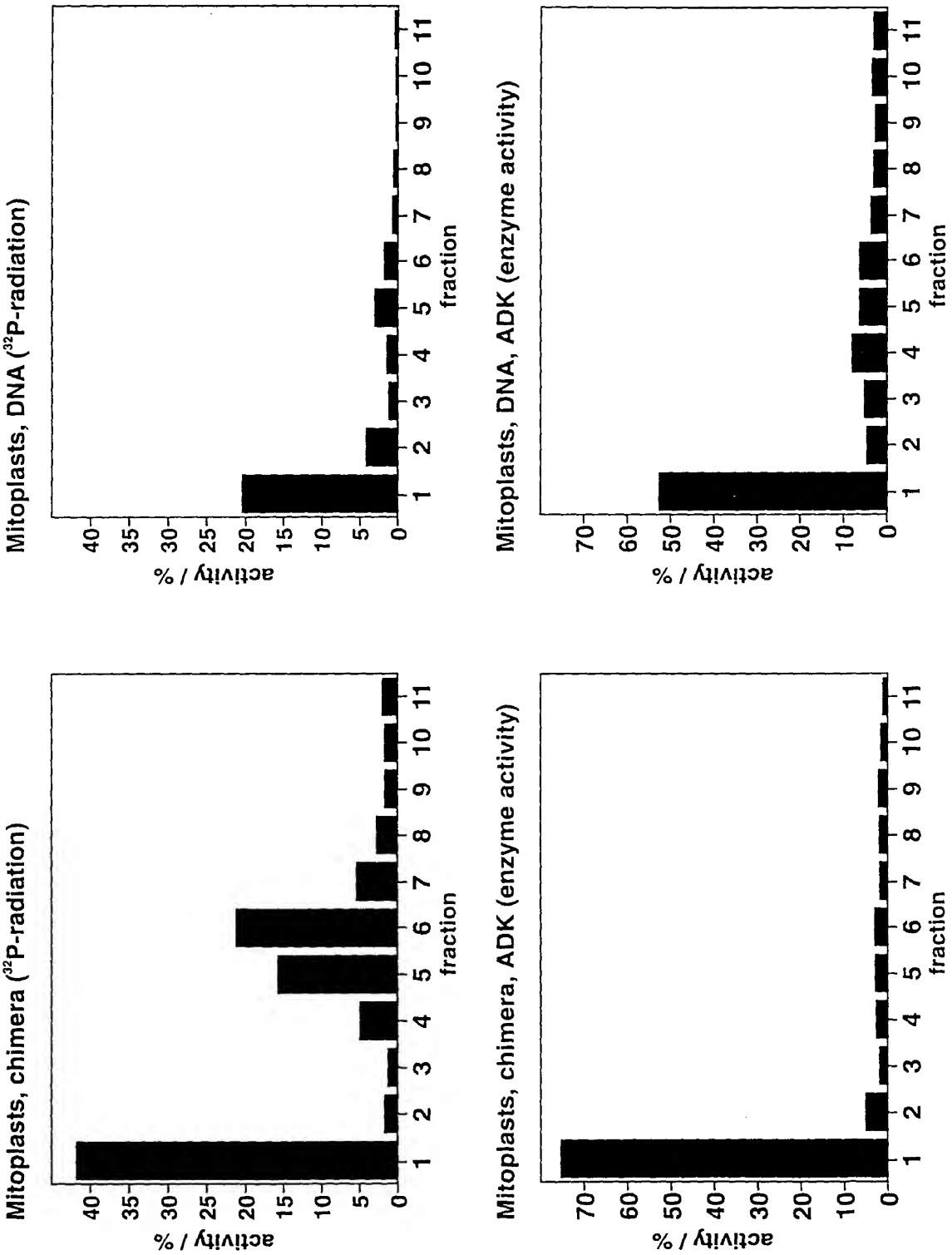

Mitoplasts, DNA (^{32}P-radiation)
Mitoplasts, DNA, ADK (enzyme activity)
Mitoplasts, chimera (^{32}P-radiation)
Mitoplasts, chimera, ADK (enzyme activity)
activity / %
fraction
0
5
10
15
20
25
30
35
40
10
20
30
40
50
60
70
1
2
3
4
5
6
7
8
9
10
11

Figure 7b

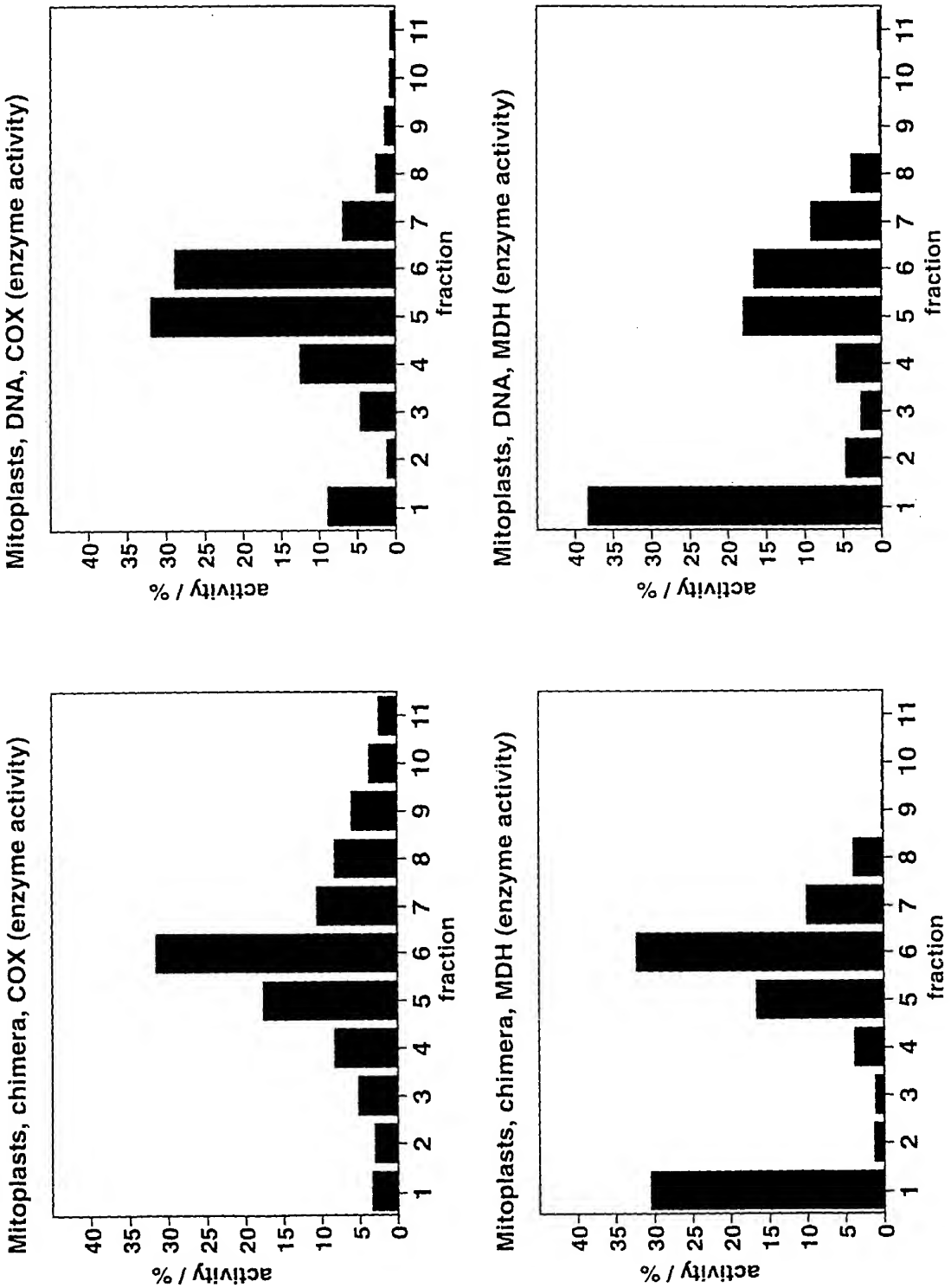

Mitoplasts, DNA, COX (enzyme activity)
Mitoplasts, DNA, MDH (enzyme activity)
Mitoplasts, chimera, COX (enzyme activity)
Mitoplasts, chimera, MDH (enzyme activity)
activity / %
fraction
0
5
10
15
20
25
30
35
40
1
2
3
4
5
6
7
8
9
10
11

Figure 8

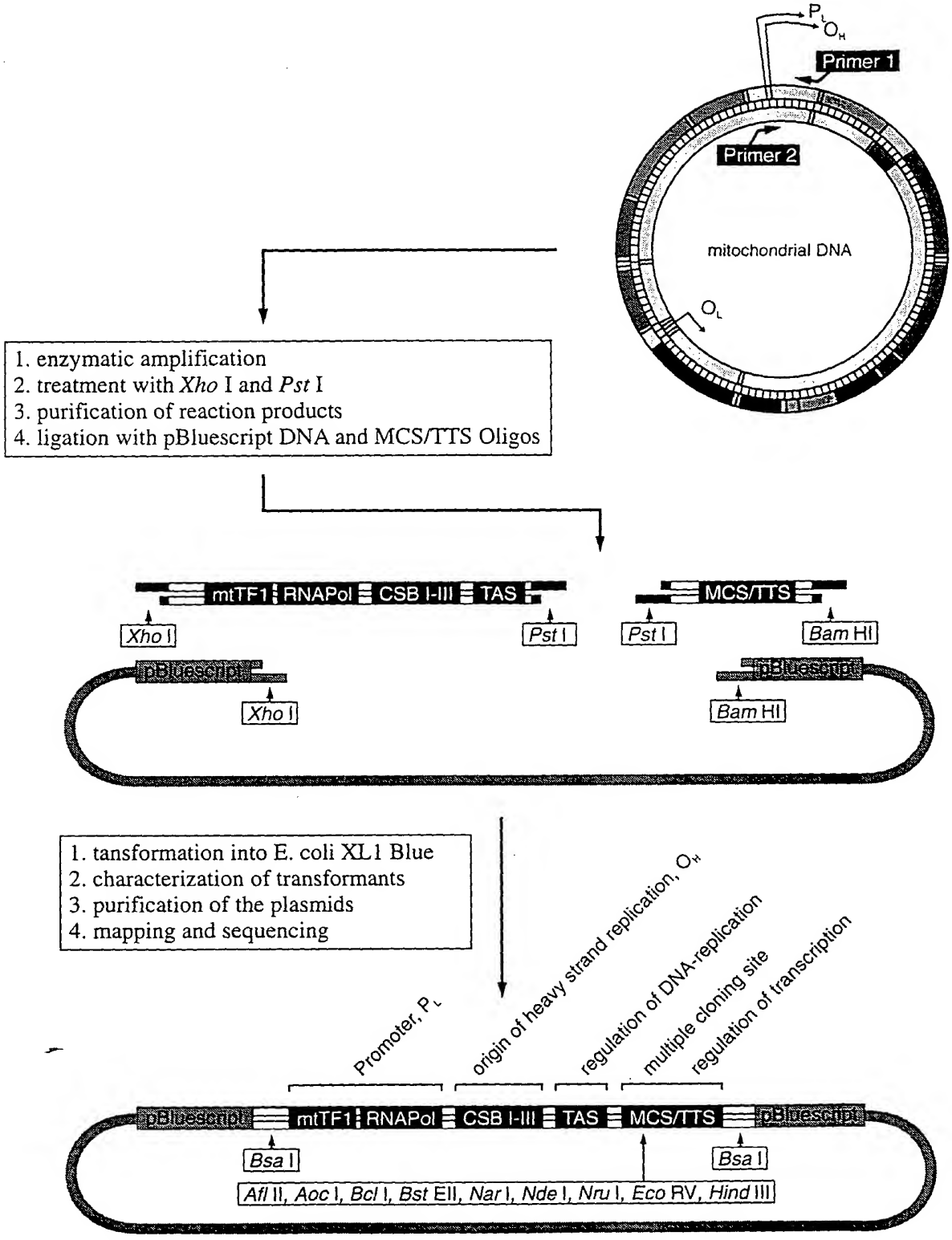

P_L
O_H
Primer 1
Primer 2
mitochondrial DNA
O_L
1. enzymatic amplification
2. treatment with Xho I and Pst I
3. purification of reaction products
4. ligation with pBluescript DNA and MCS/TTS Oligos
mtTF1
RNAPol
CSB I-III
TAS
MCS/TTS
Xho I
Pst I
Pst I
Bam HI
pBluescript
Xho I
Bam HI
1. tansformation into E. coli XL1 Blue
2. characterization of transformants
3. purification of the plasmids
4. mapping and sequencing
Promoter, P_L
origin of heavy strand replication, O_H
regulation of DNA-replication
multiple cloning site
regulation of transcription
pBluescript
mtTF1
RNAPol
CSB I-III
TAS
MCS/TTS
pBluescript
Bsa I
Bsa I
Afl II, Aoc I, Bcl I, Bst EII, Nar I, Nde I, Nru I, Eco RV, Hind III

Figure 9

Pst I
Nru I
Afl II
Aoc I
Nde I
Eco RV
Hind III
Nar I
Bst EII
Bcl I
MCS/TTS 1
Bam HI
MCS/TTS 2
Bsa I
transcription-termination-sequence, TTS
5' GATATCGCGAAGCTTAAGGCGCCTCAGGTCACCATATGATCATTTGTTAAGATGGCAGAGCCCGGTAATCGCATAAAATGAGACCG 3'
3' ACGTCTATAGCGCTTCGAATTCCGCGGAGTCCAGTGGTATACTAGTAAACAATTCTACCGTCTCGGGCCATTAGCGTATTTTACTCTGGCCTAG 5'

Figure 10


```
        10          20          30          40          50          60
CTCGAGGGTC TCAGGGGCTA ATAGAAAGGC TAGGACCAAA CCTATTTGTT TATGGGGTGA
GAGCTCCCAG AGTCCCCGAT TATCTTTCCG ATCCTGGTTT GGATAAACAA ATACCCCACT

        70          80          90          100         110         120
TGTGAGCCCG TCTAAACATT TTCAGTGTAT TGCTTTGAGG AGGTAAGCTA CATAAACTGT
ACACTCGGGC AGATTTGTAA AAGTCACATA ACGAAACTCC TCCATTCGAT GTATTTGACA

        130         140         150         160         170         180
GGGGGGTGTC TTTGGGGTTT GGTTGGTTCG GGGTATGGGG TTAGCAGCGG TGTGTGTGTG
CCCCCCACAG AAACCCCAAA CCAACCAAGC CCCATACCCC AATCGTCGCC ACACACACAC

        190         200         210         220         230         240
CTGGGTAGGA TGGGCGGGGG TTGTATTGAT GAGATTAGTA GTATGGGAGT GGGAGGGGAA
GACCCATCCT ACCCGCCCCC AACATAACTA CTCTAATCAT CATACCCTCA CCCTCCCCTT

        250         260         270         280         290         300
AATAATGTGT TAGTTGGGGG GTGACTGTTA AAAGTGCATA CCGCCAAAAG ATAAAATTTG
TTATTACACA ATCAACCCCC CACTGACAAT TTTCACGTAT GGCGGTTTTC TATTTTAAAC

        310         320         330         340         350         360
AAATCTGGTT AGGCTGGTGT TAGGGTTCTT TGTTTTTGGG GTTTGGCAGA GATGTGTTTA
TTTAGACCAA TCCGACCACA ATCCCAAGAA ACAAAAACCC CAAACCGTCT CTACACAAAT

        370         380         390         400         410         420
AGTGCTGTGG CCAGAAGCGG GGGAGGGGGG GTTTGGTGGA AATTTTTTGT TATGATGTCT
TCACGACACC GGTCTTCGCC CCCTCCCCCC CAAACCACCT TTAAAAAACA ATACTACAGA

        430         440         450         460         470         480
GTGTGGAAAG TGGCTGTGCA GACATTCAAT TGTTATTATT ATGTCCTACA AGCATTAATT
CACACCTTTC ACCGACACGT CTGTAAGTTA ACAATAATAA TACAGGATGT TCGTAATTAA

        490         500         510         520         530         540
AATTAACACA CTTTAGTAAG TATGTTCGCC TGTAATATTG AACGTAGGTG CGATAAATAA
TTAATTGTGT GAAATCATTC ATACAAGCGG ACATTATAAC TTGCATCCAC GCTATTTATT

        550         560         570         580         590         600
TAGGATGAGG CAGGAATCAA AGACAGATAC TGCGACATAG GGTGCTCCGG CTCCAGCGTC
ATCCTACTCC GTCCTTAGTT TCTGTCTATG ACGCTGTATC CCACGAGGCC GAGGTCGCAG

        610         620         630         640         650         660
TCGCAATGCT ATCGCGTGCA TACCCCCCAG ACGAAAATAC CAAATGCATG GAGAGCTCCC
AGCGTTACGA TAGCGCACGT ATGGGGGGTC TGCTTTTATG GTTTACGTAC CTCTCGAGGG

        670         680         690         700         710         720
GTGAGTGGTT AATAGGGTGA TAGACCTGTG ATCCATCGTG ATGTCTTATT TAAGGGGAAC
CACTCACCAA TTATCCCACT ATCTGGACAC TAGGTAGCAC TACAGAATAA ATTCCCCTTG

        730         740         750         760         770         780
GTGTGGGCTA TTTAGGCTTT ATGACCCTGA AGTAGGAACC AGATGTCGGA TACAGTTCAC
CACACCCGAT AAATCCGAAA TACTGGGACT TCATCCTTGG TCTACAGCCT ATGTCAAGTG
```

```
      790         800         810         820         830         840
TTTAGCTACC CCCAAGTGTT ATGGGCCCGG AGCGAGGAGA GTAGCACTCT TGTGCGGGAT
AAATCGATGG GGGTTCACAA TACCCGGGCC TCGCTCCTCT CATCGTGAGA ACACGCCCTA

      850         860         870         880         890         900
ATTGATTTCA CGGAGGATGG TGGTCAAGGG ACCCCTATCT GAGGGGGGTC ATCCATGGGG
TAACTAAAGT GCCTCCTACC ACCAGTTCCC TGGGGATAGA CTCCCCCCAG TAGGTACCCC

      910         920         930         940         950         960
ACGAGAAGGG ATTTGACTGT AATGTGCTAT GTACGGTAAA TGGCTTTATG TACTATGTAC
TGCTCTTCCC TAAACTGACA TTACACGATA CATGCCATTT ACCGAAATAC ATGATACATG

      970         980         990        1000        1010        1020
TGTTAAGGGT GGGTAGGTTT GTTGGTATCC TAGTGGGTGA GGGGTGGCTT TGGAGTTGCA
ACAATTCCCA CCCATCCAAA CAACCATAGG ATCACCCACT CCCCACCGAA ACCTCAACGT

     1030        1040        1050        1060        1070        1080
GTTGATGTGT GATAGTTGAG GGTTGATTGC TGTACTTGCT TGTAAGCATG GGGAGGGGGT
CAACTACACA CTATCAACTC CCAACTAACG ACATGAACGA ACATTCGTAC CCCTCCCCCA

     1090        1100        1110        1120        1130        1140
TTTGATGTGG ATTGGGTTTT TATGTACTAC AGGTGGTCAA GTATTTATGG TACCGTACAA
AAACTACACC TAACCCAAAA ATACATGATG TCCACCAGTT CATAAATACC ATGGCATGTT

     1150        1160        1170        1180        1190        1200
TATTCATGGT GGCTGGCAGT AATGTACGAA ATACATAGCG GTTGTTGATG GGTGAGTCAA
ATAAGTACCA CCGACCGTCA TTACATGCTT TATGTATCGC CAACAACTAC CCACTCAGTT

     1210        1220        1230        1240        1250        1260
TACTTGGGTG GTACCCAAAT CTGCTTCCCC ATGAAAGAAC AGAGAATAGT TTAAATTAGA
ATGAACCCAC CATGGGTTTA GACGAAGGGG TACTTTCTTG TCTCTTATCA AATTTAATCT

     1270        1280        1290        1300        1310        1320
ATCTTAGCTT TGGGTGCTAA TGGTGGAGTT AAAGACTTTT TCTCTGATTT GTCCTTGGAA
TAGAATCGAA ACCCACGATT ACCACCTCAA TTTCTGAAAA AGAGACTAAA CAGGAACCTT

     1330        1340        1350        1360        1370        1380
AAAGGTTTTC ATCTCCGGTT TACAAGACTG GTGTATTAGC TGCAGATATC GCGAAGCTTA
TTTCCAAAAG TAGAGGCCAA ATGTTCTGAC CACATAATCG ACGTCTATAG CGCTTCGAAT

     1390        1400        1410        1420        1430        1440
AGGCGCCTCA GGTCACCATA TGATCATTTG TTAAGATGGC AGAGCCCGGT AATCGCATAA
TCCGCGGAGT CCAGTGGTAT ACTAGTAAAC AATTCTACCG TCTCGGGCCA TTAGCGTATT

     1450
AATGAGACCG GATCC
TTACTCTGGC CTAGG
```

Figure 11

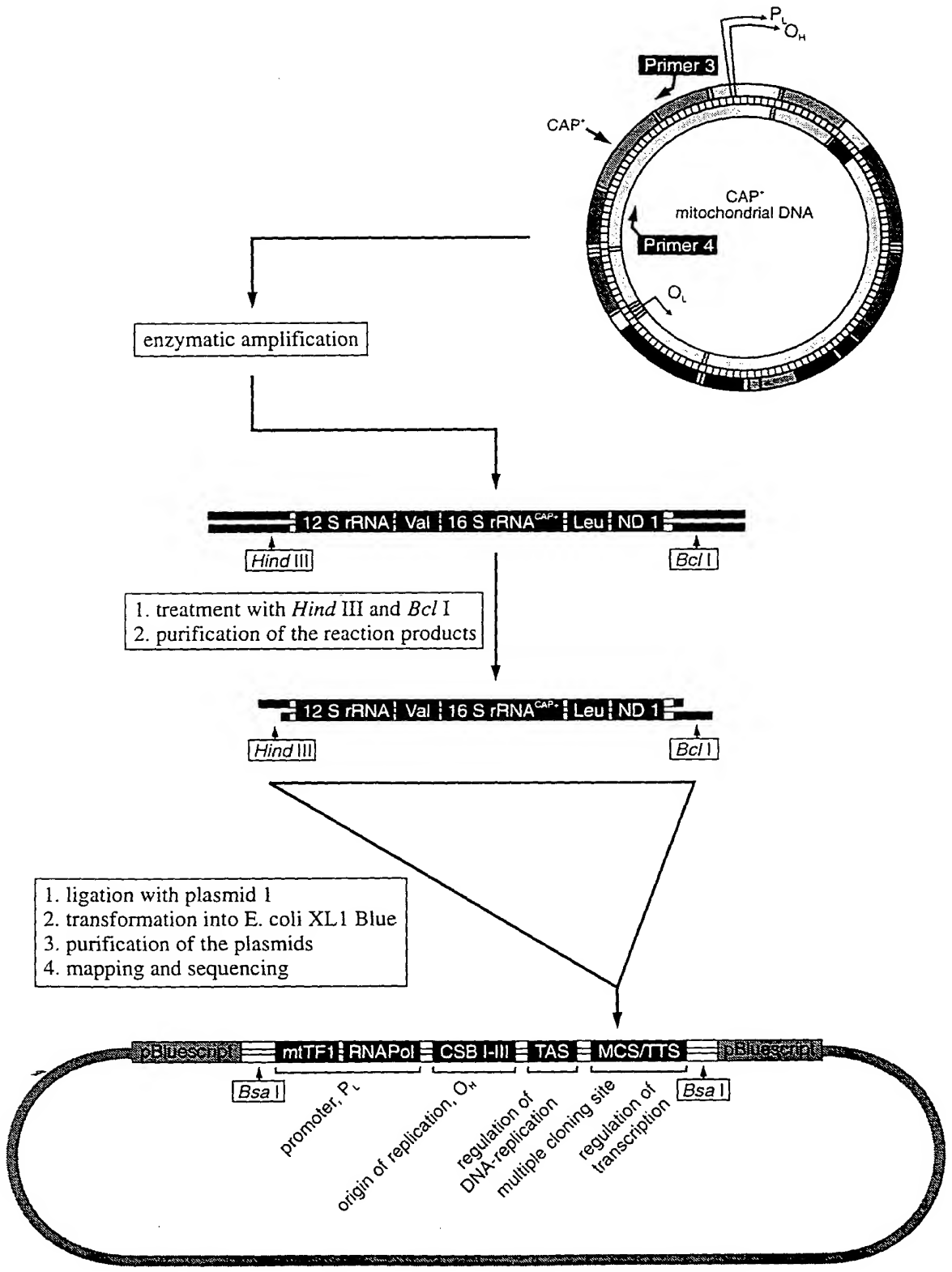

P_L
O_H
Primer 3
CAP+
CAP+ mitochondrial DNA
Primer 4
O_L
enzymatic amplification
12 S rRNA | Val | 16 S rRNA CAP+ | Leu | ND 1
Hind III
Bcl I
1. treatment with Hind III and Bcl I
2. purification of the reaction products
12 S rRNA | Val | 16 S rRNA CAP+ | Leu | ND 1
Hind III
Bcl I
1. ligation with plasmid 1
2. transformation into E. coli XL1 Blue
3. purification of the plasmids
4. mapping and sequencing
pBluescript
mtTF1
RNAPol
CSB I-III
TAS
MCS/TTS
pBluescript
Bsa I
Bsa I
promoter, P_L
origin of replication, O_H
regulation of DNA-replication
multiple cloning site
regulation of transcription

Figure 12


```
        10          20          30          40          50          60
CTCGAGGGTC TCAGGGGCTA ATAGAAAGGC TAGGACCAAA CCTATTTGTT TATGGGGTGA
GAGCTCCCAG AGTCCCCGAT TATCTTTCCG ATCCTGGTTT GGATAAACAA ATACCCCACT

        70          80          90         100         110         120
TGTGAGCCCG TCTAAACATT TTCAGTGTAT TGCTTTGAGG AGGTAAGCTA CATAAACTGT
ACACTCGGGC AGATTTGTAA AAGTCACATA ACGAAACTCC TCCATTCGAT GTATTTGACA

       130         140         150         160         170         180
GGGGGGTGTC TTTGGGGTTT GGTTGGTTCG GGGTATGGGG TTAGCAGCGG TGTGTGTGTG
CCCCCCACAG AAACCCCAAA CCAACCAAGC CCCATACCCC AATCGTCGCC ACACACACAC

       190         200         210         220         230         240
CTGGGTAGGA TGGGCGGGGG TTGTATTGAT GAGATTAGTA GTATGGGAGT GGGAGGGGAA
GACCCATCCT ACCCGCCCCC AACATAACTA CTCTAATCAT CATACCCTCA CCCTCCCCTT

       250         260         270         280         290         300
AATAATGTGT TAGTTGGGGG GTGACTGTTA AAAGTGCATA CCGCCAAAAG ATAAAATTTG
TTATTACACA ATCAACCCCC CACTGACAAT TTTCACGTAT GGCGGTTTTC TATTTTAAAC

       310         320         330         340         350         360
AAATCTGGTT AGGCTGGTGT TAGGGTTCTT TGTTTTTGGG GTTTGGCAGA GATGTGTTTA
TTTAGACCAA TCCGACCACA ATCCCAAGAA ACAAAAACCC CAAACCGTCT CTACACAAAT

       370         380         390         400         410         420
AGTGCTGTGG CCAGAAGCGG GGGAGGGGGG GTTTGGTGGA AATTTTTTGT TATGATGTCT
TCACGACACC GGTCTTCGCC CCCTCCCCCC CAAACCACCT TTAAAAAACA ATACTACAGA

       430         440         450         460         470         480
GTGTGGAAAG TGGCTGTGCA GACATTCAAT TGTTATTATT ATGTCCTACA AGCATTAATT
CACACCTTTC ACCGACACGT CTGTAAGTTA ACAATAATAA TACAGGATGT TCGTAATTAA

       490         500         510         520         530         540
AATTAACACA CTTTAGTAAG TATGTTCGCC TGTAATATTG AACGTAGGTG CGATAAATAA
TTAATTGTGT GAAATCATTC ATACAAGCGG ACATTATAAC TTGCATCCAC GCTATTTATT

       550         560         570         580         590         600
TAGGATGAGG CAGGAATCAA AGACAGATAC TGCGACATAG GGTGCTCCGG CTCCAGCGTC
ATCCTACTCC GTCCTTAGTT TCTGTCTATG ACGCTGTATC CCACGAGGCC GAGGTCGCAG

       610         620         630         640         650         660
TCGCAATGCT ATCGCGTGCA TACCCCCCAG ACGAAAATAC CAAATGCATG GAGAGCTCCC
AGCGTTACGA TAGCGCACGT ATGGGGGGTC TGCTTTTATG GTTTACGTAC CTCTCGAGGG

       670         680         690         700         710         720
GTGAGTGGTT AATAGGGTGA TAGACCTGTG ATCCATCGTG ATGTCTTATT TAAGGGGAAC
CACTCACCAA TTATCCCACT ATCTGGACAC TAGGTAGCAC TACAGAATAA ATTCCCCTTG

       730         740         750         760         770         780
GTGTGGGCTA TTTAGGCTTT ATGACCCTGA AGTAGGAACC AGATGTCGGA TACAGTTCAC
CACACCCGAT AAATCCGAAA TACTGGGACT TCATCCTTGG TCTACAGCCT ATGTCAAGTG

       790         800         810         820         830         840
TTTAGCTACC CCCAAGTGTT ATGGGCCCGG AGCGAGGAGA GTAGCACTCT TGTGCGGGAT
AAATCGATGG GGGTTCACAA TACCCGGGCC TCGCTCCTCT CATCGTGAGA ACACGCCCTA
```

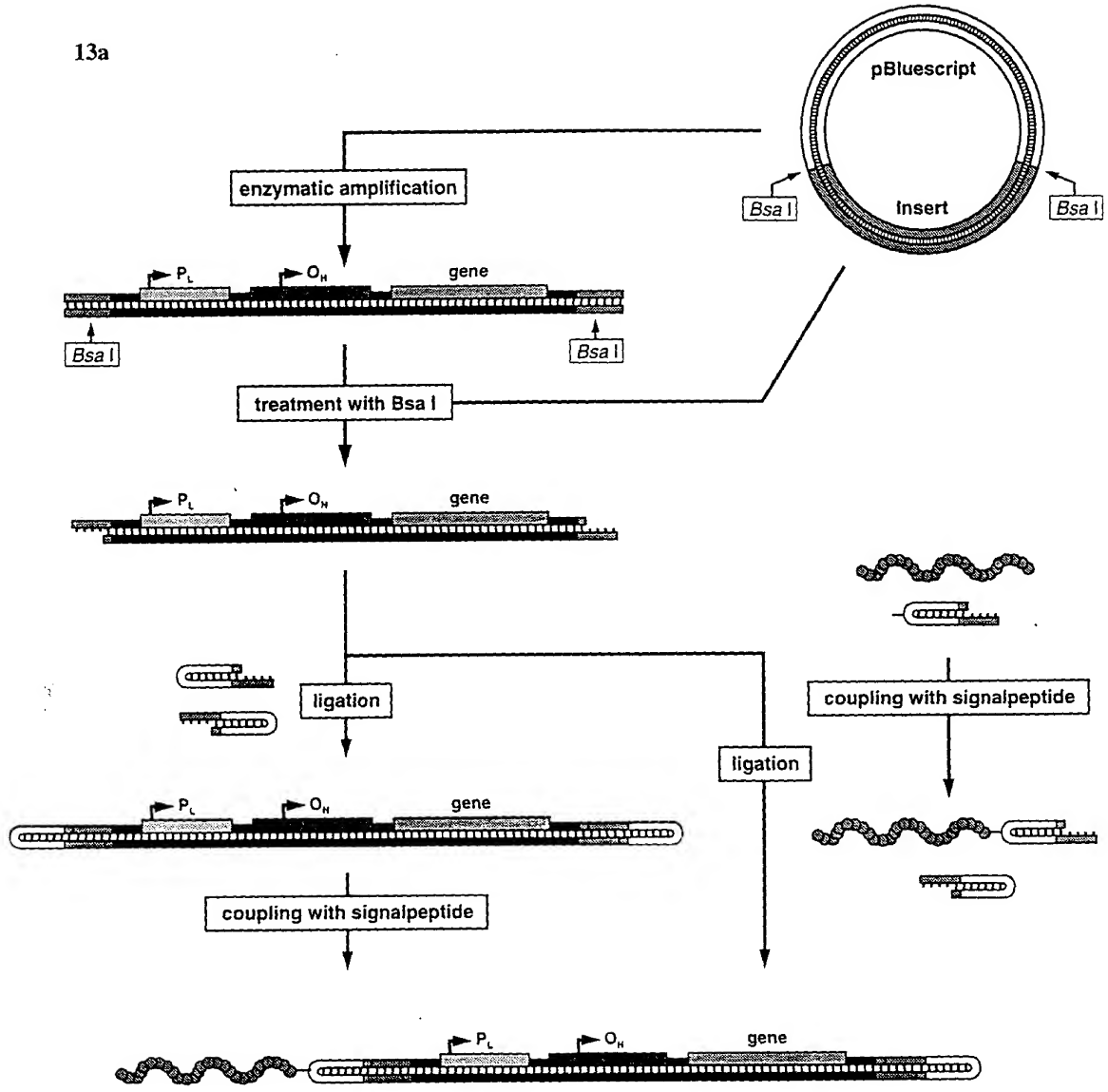
```
       850         860         870         880         890         900
ATTGATTTCA CGGAGGATGG TGGTCAAGGG ACCCCTATCT GAGGGGGGTC ATCCATGGGG
TAACTAAAGT GCCTCCTACC ACCAGTTCCC TGGGGATAGA CTCCCCCCAG TAGGTACCCC

       910         920         930         940         950         960
ACGAGAAGGG ATTTGACTGT AATGTGCTAT GTACGGTAAA TGGCTTTATG TACTATGTAC
TGCTCTTCCC TAAACTGACA TTACACGATA CATGCCATTT ACCGAAATAC ATGATACATG

       970         980         990        1000        1010        1020
TGTTAAGGGT GGGTAGGTTT GTTGGTATCC TAGTGGGTGA GGGGTGGCTT TGGAGTTGCA
ACAATTCCCA CCCATCCAAA CAACCATAGG ATCACCCACT CCCCACCGAA ACCTCAACGT

      1030        1040        1050        1060        1070        1080
GTTGATGTGT GATAGTTGAG GGTTGATTGC TGTACTTGCT TGTAAGCATG GGGAGGGGGT
CAACTACACA CTATCAACTC CCAACTAACG ACATGAACGA ACATTCGTAC CCCTCCCCCA

      1090        1100        1110        1120        1130        1140
TTTGATGTGG ATTGGGTTTT TATGTACTAC AGGTGGTCAA GTATTTATGG TACCGTACAA
AAACTACACC TAACCCAAAA ATACATGATG TCCACCAGTT CATAAATACC ATGGCATGTT

      1150        1160        1170        1180        1190        1200
TATTCATGGT GGCTGGCAGT AATGTACGAA ATACATAGCG GTTGTTGATG GGTGAGTCAA
ATAAGTACCA CCGACCGTCA TTACATGCTT TATGTATCGC CAACAACTAC CCACTCAGTT

      1210        1220        1230        1240        1250        1260
TACTTGGGTG GTACCCAAAT CTGCTTCCCC ATGAAAGAAC AGAGAATAGT TTAAATTAGA
ATGAACCCAC CATGGGTTTA GACGAAGGGG TACTTTCTTG TCTCTTATCA AATTTAATCT

      1270        1280        1290        1300        1310        1320
ATCTTAGCTT TGGGTGCTAA TGGTGGAGTT AAAGACTTTT TCTCTGATTT GTCCTTGGAA
TAGAATCGAA ACCCACGATT ACCACCTCAA TTTCTGAAAA AGAGACTAAA CAGGAACCTT

      1330        1340        1350        1360        1370        1380
AAAGGTTTTC ATCTCCGGTT TACAAGACTG GTGTATTAGC TGCAGATATC GCGAAGCTTG
TTTCCAAAAG TAGAGGCCAA ATGTTCTGAC CACATAATCG ACGTCTATAG CGCTTCGAAC

      1390        1400        1410        1420        1430        1440
TAACATGGTA AGTGTACTGG AAAGTGCACT TGGACGAACC AGAGTGTAGC TTAACACAAA
ATTGTACCAT TCACATGACC TTTCACGTGA ACCTGCTTGG TCTCACATCG AATTGTGTTT

      1450        1460        1470        1480        1490        1500
GCACCCAACT TACACTTAGG AGATTTCAAC TTAACTTGAC CGCTCTGAGC TAAACCTAGC
CGTGGGTTGA ATGTGAATCC TCTAAAGTTG AATTGAACTG GCGAGACTCG ATTTGGATCG

      1510        1520        1530        1540        1550        1560
CCCAAACCCA CTCCACCTTA CTACCAGACA ACCTTAGCCA AACCATTTAC CCAAATAAAG
GGGTTTGGGT GAGGTGGAAT GATGGTCTGT TGGAATCGGT TTGGTAAATG GGTTTATTTC

      1570        1580        1590        1600        1610        1620
TATAGGCGAT AGAAATTGAA ACCTGGCGCA ATAGATATAG TACCGCAAGG GAAAGATGAA
ATATCCGCTA TCTTTAACTT TGGACCGCGT TATCTATATC ATGGCGTTCC CTTTCTACTT

      1630        1640        1650        1660        1670        1680
AAATTATAAC CAAGCATAAT ATAGCAAGGA CTAACCCCTA TACCTTCTGC ATAATGAATT
TTTAATATTG GTTCGTATTA TATCGTTCCT GATTGGGGAT ATGGAAGACG TATTACTTAA
```

```
      1690        1700        1710        1720        1730        1740
AACTAGAAAT AACTTTGCAA GGAGAGCCAA AGCTAAGACC CCCGAAACCA GACGAGCTAC
TTGATCTTTA TTGAAACGTT CCTCTCGGTT TCGATTCTGG GGGCTTTGGT CTGCTCGATG

      1750        1760        1770        1780        1790        1800
CTAAGAACAG CTAAAAGAGC ACACCCGTCT ATGTAGCAAA ATAGTGGGAA GATTTATAGG
GATTCTTGTC GATTTTCTCG TGTGGGCAGA TACATCGTTT TATCACCCTT CTAAATATCC

      1810        1820        1830        1840        1850        1860
TAGAGGCGAC AAACCTACCG AGCCTGGTGA TAGCTGGTTG TCCAAGATAG AATCTTAGTT
ATCTCCGCTG TTTGGATGGC TCGGACCACT ATCGACCAAC AGGTTCTATC TTAGAATCAA

      1870        1880        1890        1900        1910        1920
CAACTTTAAA TTTGCCCACA GAACCCTCTA AATCCCCTTG TAAATTTAAC TGTTAGTCCA
GTTGAAATTT AAACGGGTGT CTTGGGAGAT TTAGGGGAAC ATTTAAATTG ACAATCAGGT

      1930        1940        1950        1960        1970        1980
AAGAGGAACA GCTCTTTGGA CACTAGGAAA AAACCTTGTA GAGAGAGTAA AAAATTTAAC
TTCTCCTTGT CGAGAAACCT GTGATCCTTT TTTGGAACAT CTCTCTCATT TTTTAAATTG

      1990        2000        2010        2020        2030        2040
ACCCATAGTA GGCCTAAAAG CAGCCACCAA TTAAGAAAGC GTTCAAGCTC AACACCCACT
TGGGTATCAT CCGGATTTTC GTCGGTGGTT AATTCTTTCG CAAGTTCGAG TTGTGGGTGA

      2050        2060        2070        2080        2090        2100
ACCTAAAAAA TCCCAAACAT ATAACTGAAC TCCTCACACC CAATTGGACC AATCTATCAC
TGGATTTTTT AGGGTTTGTA TATTGACTTG AGGAGTGTGG GTTAACCTGG TTAGATAGTG

      2110        2120        2130        2140        2150        2160
CCTATAGAAG AACTAATGTT AGTATAAGTA ACATGAAAAC ATTCTCCTCC GCATAAGCCT
GGATATCTTC TTGATTACAA TCATATTCAT TGTACTTTTG TAAGAGGAGG CGTATTCGGA

      2170        2180        2190        2200        2210        2220
GCGTCAGATT AAAACACTGA ACTGACAATT AACAGCCCAA TATCTACAAT CAACCAACAA
CGCAGTCTAA TTTTGTGACT TGACTGTTAA TTGTCGGGTT ATAGATGTTA GTTGGTTGTT

      2230        2240        2250        2260        2270        2280
GTCATTATTA CCCTCACTGT CAACCCAACA CAGGCATGCT CATAAGGAAA GGTTAAAAAA
CAGTAATAAT GGGAGTGACA GTTGGGTTGT GTCCGTACGA GTATTCCTTT CCAATTTTTT

      2290        2300        2310        2320        2330        2340
AGTAAAAGGA ACTCGGCAAA TCTTACCCCG CCTGTTTACC AAAAACATCA CCTCTAGCAT
TCATTTTCCT TGAGCCGTTT AGAATGGGGC GGACAAATGG TTTTTGTAGT GGAGATCGTA

      2350        2360        2370        2380        2390        2400
CACCAGTATT AGAGGCACCG CCTGCCCAGT GACACATGTT TAACGGCCGC GGTACCCTAA
GTGGTCATAA TCTCCGTGGC GGACGGGTCA CTGTGTACAA ATTGCCGGCG CCATGGGATT

      2410        2420        2430        2440        2450        2460
CCGTGCAAAG GTAGCATAAT CACTTGTTCC TTAAATAGGG ACCTGTATGA ATGGCTCCAC
GGCACGTTTC CATCGTATTA GTGAACAAGG AATTTATCCC TGGACATACT TACCGAGGTG

      2470        2480        2490        2500        2510        2520
GAGGGTTCAG CTGTCTCTTA CTTTTAACCA GTGAAATTGA CCTGCCCGTG AAGAGGCGGG
CTCCCAAGTC GACAGAGAAT GAAAATTGGT CACTTTAACT GGACGGGCAC TTCTCCGCCC
```

```
      2530        2540        2550        2560        2570        2580
CATAACACAG CAAGACGAGA AGACCCTATG GAGCTTTAAT TTATTAATGC AAACAGTACC
GTATTGTGTC GTTCTGCTCT TCTGGGATAC CTCGAAATTA AATAATTACG TTTGTCATGG

      2590        2600        2610        2620        2630        2640
TAACAAACCC ACAGGTCCTA AACTACCAAA CCTGCATTAA AAATTTCGGT TGGGGCGACC
ATTGTTTGGG TGTCCAGGAT TTGATGGTTT GGACGTAATT TTTAAAGCCA ACCCCGCTGG

      2650        2660        2670        2680        2690        2700
TCGGAGCAGA ACCCAACCTC CGAGCAGTAC ATGCTAAGAC TTCACCAGTC AAAGCGAACT
AGCCTCGTCT TGGGTTGGAG GCTCGTCATG TACGATTCTG AAGTGGTCAG TTTCGCTTGA

      2710        2720        2730        2740        2750        2760
ACTATACTCA ATTGATCCAA TAACTTGACC AACGGAACAA GTTACCCTAG GGATAACAGC
TGATATGAGT TAACTAGGTT ATTGAACTGG TTGCCTTGTT CAATGGGATC CCTATTGTCG

      2770        2780        2790        2800        2810        2820
GCAATCCTAT TCTAGAGTCC ATATCAACAA TAGGGTTTAC GACCTCGATG TTGGATCAGG
CGTTAGGATA AGATCTCAGG TATAGTTGTT ATCCCAAATG CTGGAGCTAC AACCTAGTCC

      2830        2840        2850        2860        2870        2880
ACATCCCGAT GGTGCAGCCG CTATTAAAGG TTCGTTTGTT CAACGATTAA AGTCCTACGT
TGTAGGGCTA CCACGTCGGC GATAATTTCC AAGCAAACAA GTTGCTAATT TCAGGATGCA

      2890        2900        2910        2920        2930        2940
GATCTGAGTT CAGACCGGAG TAATCCAGGT CGGTTTCTAT CTACCTTCAA ATTCCTCCCT
CTAGACTCAA GTCTGGCCTC ATTAGGTCCA GCCAAAGATA GATGGAAGTT TAAGGAGGGA

      2950        2960        2970        2980        2990        3000
GTACGAAAGG ACAAGAGAAA TAAGGCCTAC TTCACAAAGC GCCTTCCCCC GTAAATGATA
CATGCTTTCC TGTTCTCTTT ATTCCGGATG AAGTGTTTCG CGGAAGGGGG CATTTACTAT

      3010        3020        3030        3040        3050        3060
TCATCTCAAC TTAGTATTAT ACCCACACCC ACCCAAGAAC AGGGTTTGTT AAGATGGCAG
AGTAGAGTTG AATCATAATA TGGGTGTGGG TGGGTTCTTG TCCCAAACAA TTCTACCGTC

      3070        3080        3090        3100        3110        3120
AGCCCGGTAA TCGCATAAAA CTTAAAACTT TACAGTCAGA GGTTCAATTC CTCTTCTTAA
TCGGGCCATT AGCGTATTTT GAATTTTGAA ATGTCAGTCT CCAAGTTAAG GAGAAGAATT

      3130        3140        3150        3160        3170        3180
CAACATACCC ATGGCCAACC TCCTACTCCT CATTGTACCC ATTCTAATCG CAATGGCTGA
GTTGTATGGG TACCGGTTGG AGGATGAGGA GTAACATGGG TAAGATTAGC GTTACCGACT

      3190        3200        3210        3220        3230
TCATTTGTTA AGATGGCAGA GCCCGGTAAT CGCATAAAAT GAGACCGGAT CC
AGTAAACAAT TCTACCGTCT CGGGCCATTA GCGTATTTTA CTCTGGCCTA GG
```

Figure 13

13a



13b

CCCCGGGTACCTTGCGAGCCC$_X$
CCCATGGAACGCTCGGG

HP 1 (X=modified dT)

TTTTGCAGCTGGATCCCGGGC$_A$
CGTCGACCTAGGGCCCG

HP 2

Figure 14

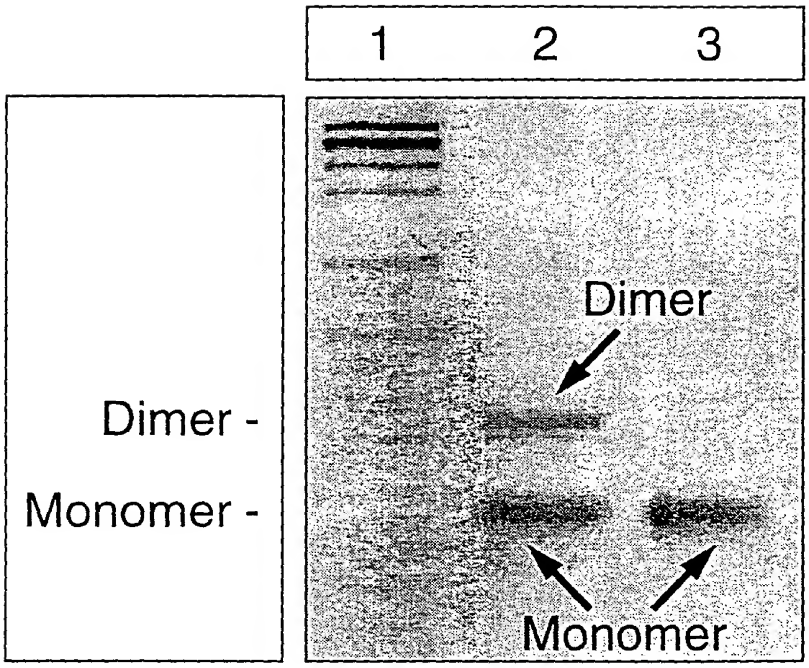
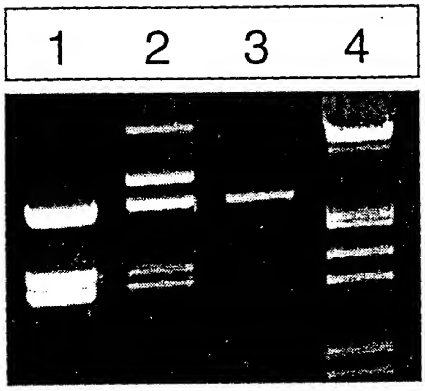


Figure 15

15a



15b

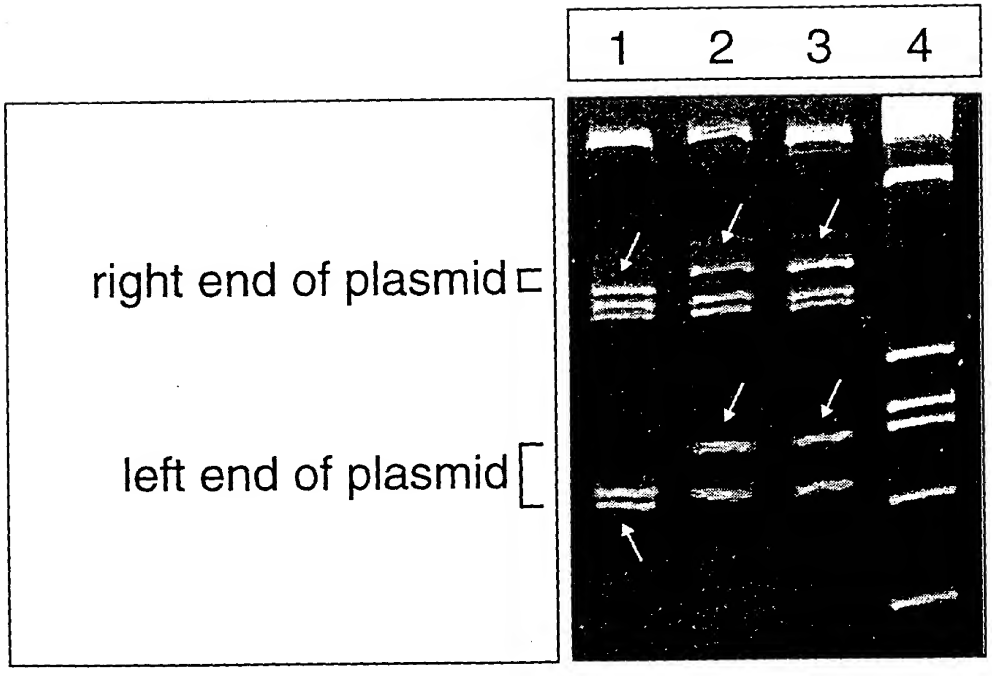


Figure 16